

ロー・パワー・デザイン 特別企画 第2弾

第2章

デバイス設計の視点で見た低消費電力技術

―― アーキテクチャからレイアウトまで，あらゆる設計階層で電力を削る

宇佐美 公良

集積回路の設計段階で消費電力を低減する技術は，LSI開発のカギを握る技術として重要性を増している．ここでは，LSI設計における低消費電力問題について解説する．LSIの消費電力は，ダイナミック電力，ショート・サーキット電力，リーク電力からなる．このうちダイナミック電力とリーク電力を低減する手法について，順に説明していく．（編集部）

情報機器の主要部品である集積回路において，低消費電力化へのニーズが高まっています．この背景には，二つの異なる大きな市場からの要求があります．一つは，携帯型情報機器の市場です．電池駆動の機器なので，電池の充電を頻繁に行わなくても済むように，電子部品にはできるだけ消費電力の小さいものが要求されます．一方，最近の携帯電話では，通常の電話機能だけでなく，映像を見る，音楽を聴く，ゲームを楽しむといった豊富な機能が搭載されており，これらを実現するには，集積回路に従来以上の高性能が要求されます．一般に，性能を引き上げようとすると消費電力は大きくなるので，これまで以上に低消費電力化への注力が必要となります．

集積回路の低消費電力化を要求するもう一つの市場が，ハイエンドのサーバ，あるいはデスクトップ・パソコンの市場です．これらは電池駆動ではないので，別の理由で低消費電力化が要求されています．それは，チップの発熱です．動作周波数がGHzのオーダで，内部で高度な並列処理を行うCPUでは，今や1チップで100Wを超えるものが現われています．このチップがどれくらい熱くなるかを想像するために，ホットプレートを例にとって比べてみます．チップの大きさはだいたい1cm四方と考えて，この面積でホットプレートが消費する電力を試算すると10W程度です．ということは，100Wのチップはホットプレートの10倍もの電力を消費し，発熱することになります．実際，しばらく前のCPUチップでしたが，動作中に放熱用の器材を取り外した瞬間，自分自身の発熱でチップの温度が上がり，白煙を上げて焼け焦げてしまうということも起こりました．放熱対策のために，機器のコストが上昇するという状況を招いています．

こういった背景から，集積回路の設計段階で消費電力を低減する技術は，LSI開発のカギを握る技術として重要性を増しています．

● まず，集積回路の電力消費のメカニズムを理解する

低消費電力技術について述べる前に，集積回路の電力はどのようなメカニズムで消費されるのかについて，簡単に触れたいと思います．CMOSのLSIでは，消費電力は大きく三つの成分からなります．それはダイナミック電力P_D，ショート・サーキット電力P_{SC}，リーク電力P_Lであり，消費電力はこれらの和となります．まず，ダイナミック電力ですが，これはCMOS回路の出力が「‘0’から‘1’」または「‘1’から‘0’」へ遷移するときに，負荷容量の充放電動作によって消費される電力です．**図1**に示すインバータを例にとると，インバータの出力が「‘0’から‘1’」に遷移するとき，電源V_{DD}からpMOSを通って負荷容量Cへ充電電流が流れます．pMOSにはON抵抗があるので，この中を通って充電電流が流れる際にpMOSでジュール熱が発生し，

KeyWord
CMOS，ダイナミック電力，ショート・サーキット電力，リーク電力，スイッチング確率，GIDL電流，低消費電力論理合成，ゲーテッド・クロック，オペランド・アイソレーション，Voltege Island手法

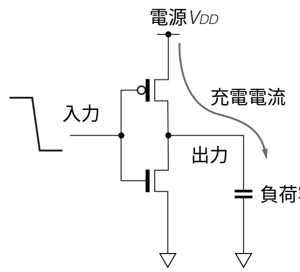


図1　CMOSインバータ回路におけるスイッチング動作と電力消費

インバータの入力が‘1’から‘0’に切り替わるとき，出力は‘0’から‘1’に遷移する．このとき，電源V_{DD}からpMOSを通って電流が流れ，負荷容量Cを充電する．pMOSにはON抵抗があるので，充電電流によりpMOSでジュール熱が発生し，電力消費が起こる．出力が‘1’から‘0’に遷移するときは，nMOSを通ってグラウンドへ放電電流が流れるので，nMOSで電力消費が生じる．

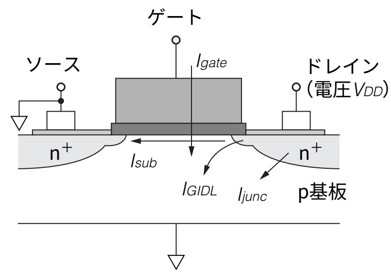


図2　nMOSトランジスタの断面構造とリーク電流

トランジスタがOFF状態でソース・ドレイン間に流れるサブスレッショルド・リーク電流I_{sub}，ゲート絶縁膜を通してトンネル効果により流れるゲート・リーク電流I_{gate}，逆バイアスが印加されたPN接合でのジャンクション・リーク電流I_{junc}，ゲート電極下のドレイン端に高電界がかかることによりドレインから基板へ流れるGIDL（gate induced drain leakage）電流I_{GIDL}が，おもなリーク電流の成分である．

電力消費が起こります．一方，インバータの出力が「‘1’から‘0’」に遷移するときには，nMOSを通ってグラウンドへ放電電流が流れるので，その際にnMOSで電力消費が生じます．このようにダイナミック電力P_Dは，出力負荷の充放電電流によって消費する電力であり，次式で表されます．

$$P_D = CV^2 f\alpha \quad \cdots\cdots(1)$$

ここで，Cは負荷容量，Vは電源電圧，fは動作周波数，αはスイッチング確率です．スイッチング確率とは，ゲートの出力が遷移する確率のことです．論理回路網を考えたときに，それぞれのゲートはかならずしも毎サイクル遷移するわけではありません．入力が切り換わらなければ出力は切り換わりませんし，論理によっては入力が切り換わっても出力が切り換わらない場合もあります．例えばANDゲートは，入力の一つが‘0’ならば，ほかの入力が切り換わっても出力は‘0’のまま遷移しません．ダイナミック電力は，出力が遷移するときだけ消費するので，各ゲートの消費電力を求める際には，それぞれのスイッチング確率を考慮する必要があります．

一方，スイッチングの過程でpMOSとnMOSが同時にONになる短い期間がありますが，この短い期間，電源からグラウンドへ貫通電流が流れます．これによって消費する電力がショート・サーキット電力P_{SC}です．

リーク電力P_Lは，ゲートがスイッチングしないときにも消費する電力で，**図2**に示すように，MOSトランジスタおよび周辺で流れる「漏れ電流（リーク電流）」によって生じます．MOSトランジスタがOFF状態でソース・ドレイン間に流れるサブスレッショルド・リーク電流I_{sub}や，薄いゲート絶縁膜を通してトンネル効果によって流れるゲート・リーク電流I_{gate}が代表的です．このほか，逆バイアスが印加されたPN接合におけるジャンクション・リーク電流I_{junc}や，ゲート電極下のドレイン端に高い電界がかかることによりドレインから基板へ流れるGIDL（gate induced drain leakage）電流I_{GIDL}も存在します．なお，製造プロセスの微細化とともに，現在ではサブスレッショルド・リーク電流とゲート・リーク電流が支配的になりつつあります．

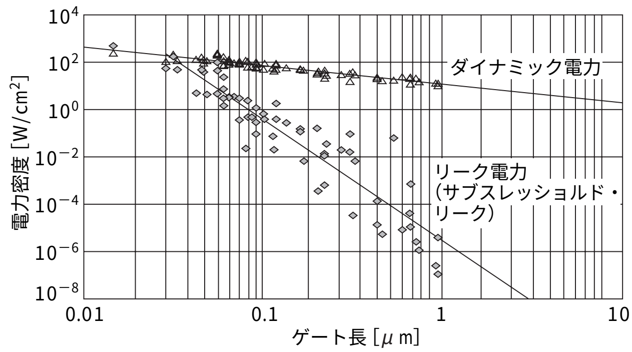


図3　集積回路の消費電力のトレンド(1)

横軸にゲート長，縦軸にチップ1cm^{2}当たりの消費電力をプロットしたもので，ゲート長が小さくなるにつれてダイナミック電力，リーク電力は，いずれも増加の一途をたどっている．とくに，プロセスの微細化とともにリーク電力が指数関数的に増えており，今後，消費電力全体の中でリーク電力の占める割合が増大する．

消費電力全体の中で，通常，もっとも大きな成分はダイナミック電力ですが，**図3**に示すように，今後プロセスの微細化が進むにつれて，リーク電力の占める割合が増大します．これを踏まえ，ダイナミック電力とリーク電力に焦点を当てて，低消費電力技術を紹介していきます．



1 ダイナミック電力を低減する代表的な技術

集積回路の低消費電力技術には代表的なものがいくつかありますが，それらの詳細を述べる前に，まず特徴によって分類し，全体の姿をとらえてみたいと思います．

ダイナミック電力を低減する技術では，$CV^2f\alpha$の値を設計段階でいかに小さくできるかが勝負となります．方法としては，C，V，f，αという四つの物理量のうち，どれをターゲットにするのかによって，**図4**のように分類できます．これらは「Cとαの積($C \times \alpha$)を低減する技術」，「Vを低減する技術」，「fを低減する技術」に大別できます．

$C \times \alpha$を低減する技術では，Cの大きい部分でαを極力低減したり，逆に，αの大きい部分に注目してCを低減したりといった手法がとられます．この技術はさらに二つに分類でき，「$C \times \alpha$の積を平均的に最小化する技術」と「Cの大きい部分でスイッチング動作を選択的に停止してαを小さくする技術」に分けることができます．前者の例は，低消費電力論理合成です．また，後者の例として，ゲーテッド・クロックとオペランド・アイソレーションが挙げられます．ゲーテッド・クロックは，クロックのスイッチングを選択的に停止してダイナミック電力を低減する技術です．オペランド・アイソレーションでは，演算器を選択的に停止することにより，電力を低減します．

一方，Vあるいはfを低減する技術では，電力低減と同時に，性能(動作速度)の維持が大きな課題となります．Vあるいはfを下げると性能が低下してしまうためです．これを解決する手法として，複数の電源電圧をLSI内で使用するマルチV_{DD}技術，および並列化の利用による低電力技術があります．

以下では，ダイナミック電力を低減する技術について，順に説明していきます(リーク電力を低減する技術については後述)．

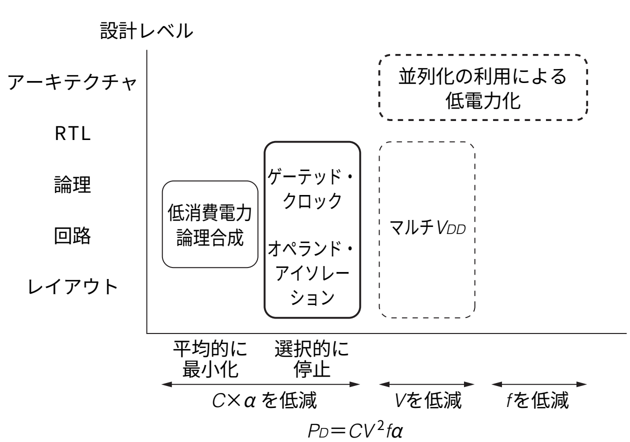


図4　ダイナミック電力低減技術の分類

ダイナミック電力を低減する技術は，$C \times \alpha$を低減する技術，Vを低減する技術，fを低減する技術に大別できる．

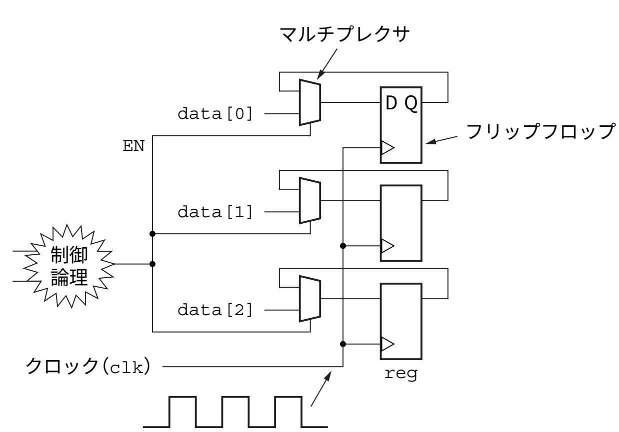
● ゲーテッド・クロック：不要なクロックのトグルを停止

現在，大半のシステムLSIはクロック同期式回路で実現されており，LSI内部に多数存在するフリップフロップ(以降FF)にはクロック信号が接続されています．このため，クロック信号には全体として，非常に大きな負荷容量が付いていることになります．さらに，ほかの論理信号と違って，クロック信号は毎サイクルかならずトグルするので，スイッチング確率も大きいことになります．クロック信号のトグルに伴って消費されるダイナミック電力は，LSI内部でも大きな割合を占めるため，低消費電力化の際，初めに検討すべき部分はクロックです．

クロックのダイナミック電力を低減する手法で，もっとも代表的なのが「ゲーテッド・クロック」です．ゲーテッ

```
always @(posedge clk)
   if(EN==1)
     reg <= data;
```

(**a**) 順序回路を表現したRTL記述の例(Verilog HDL)



(**b**) 論理合成によって生成される回路

図5　ゲーテッド・クロックの背景

(a)のRTL記述(レジスタへの代入動作)に対して，市販の論理合成ツールでは(b)に示すような回路が生成される．(b)の回路では，フリップフロップのデータを保持する場合にもクロック信号clkがトグルし，ダイナミック電力をむだに消費してしまう．

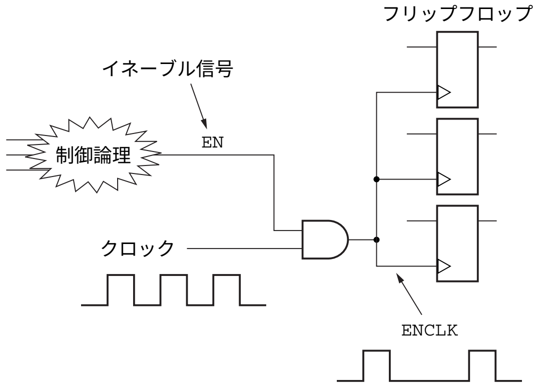


図6　ゲーテッド・クロックを施した回路

図5の例に対してゲーテッド・クロックを施した回路．クロック信号にANDゲートが挿入されている．イネーブル信号ENが'0'のとき（フリップフロップのデータを保持する場合），フリップフロップへのクロック信号ENCLKのトグルが抑えられるため，むだなダイナミック電力を消費しない．

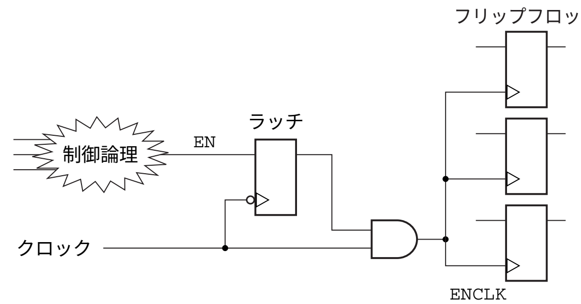


図7　ゲーテッド・クロックの実装手法（クロック信号へのグリッチ伝搬の防止）

ゲーテッド・クロックでは，クロック信号が'1'の期間にイネーブル信号ENにグリッチが発生すると，フリップフロップのクロック端子に伝搬して誤動作を引き起こす．これを防止するため，EN信号にラッチを挿入し，クロックが'1'の期間に値を保持する構成にする．

ド・クロックはクロック・ゲーティングとも呼ばれ，クロックにANDゲートを挿入して，クロックのトグルが不要な場合にトグルを止めてしまう方法です．

クロックのトグルが不要な場合とは，どのような場合でしょう．それを**図5**に示す例で説明します．**図5(a)**は，簡単な順序回路を表現したVerilog HDLのRTL記述の例で，イネーブル信号`EN`が'1'ならば，クロック`clk`の立ち上がりでデータ`data`をレジスタ`reg`に取り込む，という記述です．このRTL記述を市販の論理合成ツールに入力すると，**図5(b)**のような回路が生成されます．FFの入力にマルチプレクサが置かれ，`EN`が'1'の場合は`data`信号を選択してFFのD入力に供給します．そしてこの値を，`clk`の立ち上がりでFFに取り込みます．一方，`EN`が'0'の場合は，FFのQ出力がマルチプレクサを介してFFに戻され，結果的にFFのデータが保持されることになります．なお，マルチプレクサと等価的な論理が，ANDやORなどの論理ゲートで構成される場合もあります．

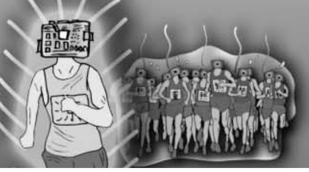
図5(b)の回路では，FFのデータを保持する場合にもクロック信号`clk`はトグルするのですが，実はこのトグルは不要なのです．理由は，通常のFF回路は電源を切らない限り，クロックを止めてもデータを保持できるからです．つまり，データを保持する期間，クロックをトグルすることは，単にダイナミック電力をむだに消費していることになります．ゲーテッド・クロックを使うと，この期間，クロックのトグルを停止することができます．

先ほどの例にゲーテッド・クロックを施すと**図6**のようになります．マルチプレクサが削除され，クロック信号にANDゲートが挿入された構成になっています．クロック信号はイネーブル信号`EN`とANDがとられている点が特徴です．`EN`が'1'のときはクロック信号がANDゲートを介して`ENCLK`として出力され，FFに伝達されるため，`ENCLK`の立ち上がりで`data`がFFに格納されます．一方，`EN`が'0'のときはクロック信号の値にかかわらず，ANDゲートの出力は'0'となるため，`ENCLK`は'0'で固定され，FFは値を保持することになります．

このようにゲーテッド・クロックでは，FFへのクロック信号に対して，トグルの必要がないときにトグルを停止するので，ダイナミック電力を低減できます．なお，クロックのトグルに伴ってFF回路の内部でもダイナミック電力を消費します．ゲーテッド・クロックにより，FF回路の内部のダイナミック電力も低減します．

ゲーテッド・クロックを実装する際に，注意するべきことがあります．それは，クロック信号へのグリッチ伝搬です．**図6**の回路では，クロック信号が'1'の期間にイネーブル信号`EN`にグリッチが発生すると，それがそのままFFのクロック端子に伝搬して誤動作を引き起こしてしまいます．これを防止するため，**図7**のように`EN`信号にラッチを挿入し，クロックが'1'の期間には値を保持する構成にします．このことから，このラッチはデグリッチ・ラッチと呼ばれます．

現在，市販のEDAツールでは，**図5(a)**に示すRTL記述から自動的に**図7**の回路を合成する，ゲーテッド・クロックの自動生成手法が実用化されています．なお，RTL記述の中にイネーブル信号が明示的に記されていない場合でも，ゲーテッド・クロックの自動生成が可能な場合があり



ます．RTL 記述の内容から，レジスタの値を保持するか否かの制御論理を抽出できる場合には，ゲーテッド・クロックの自動生成が行えます．

● オペランド・アイソレーション：演算器の動作を停止

ビット幅の大きい加算器や乗算器が多数搭載されるようになると，これらの演算器の消費電力を無視できなくなります．演算器は一般に，組み合わせ論理回路で構成されているので，入力信号が変化すると演算器内部でスイッチングが生じます．演算器の入力信号を‘0’または‘1’の値に固定することで，演算器内部のスイッチングを停止させ，ダイナミック電力を低減する手法がオペランド・アイソレーションです．

コンピュータ・アーキテクチャの分野では，演算の対象となるデータをオペランドといいますが，この手法は入力データを分離し固定してしまうことで，演算器の動作を停止する手法といえます．では，どのような場合に，入力データを固定できるのでしょうか．

図8(a)は，演算器（ここでは乗算器）とマルチプレクサが接続された構造を示しており，乗算結果Pとデータ A のいずれか一方のデータがマルチプレクサで選択され，Z として出力されます．ここで，データ A が選択される場合を考えます．この場合，P の値は使われないので，乗算器内部での演算，およびそれに伴うスイッチングは，まったく無意味なものとなります．したがって，この場合をなんらかの方法で検知して，乗算器の入力データを固定してしまえばよいわけです．検知するためのもっともシンプルな方法は，マルチプレクサの制御信号を使う方法です．マルチプレクサの制御信号 S が‘1’のとき，乗算結果 P を Z に出力し，S が‘0’のときにデータ A を Z に出力するのなら，S が‘0’のときに乗算結果は使われないので，乗算器の入力データを固定できます．

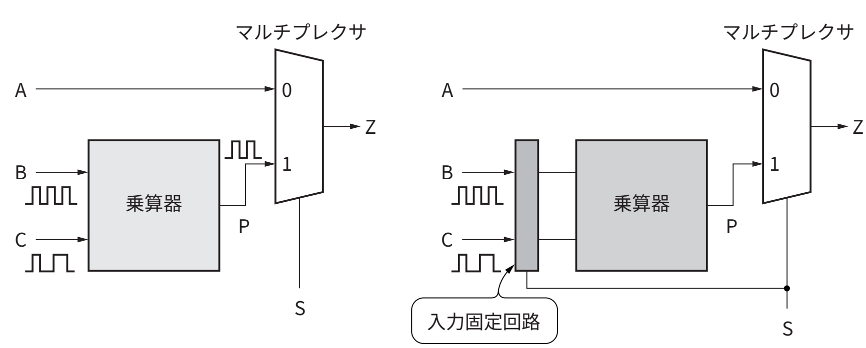
乗算器にオペランド・アイソレーションを適用した例を**図8(b)**に示します．乗算器の入力データを固定する回路（入力固定回路）が新たに追加され，信号 S で制御されます．入力固定回路は AND ゲート，OR ゲート，ラッチのいずれを用いても実現できますが，AND ゲートまたは OR ゲートを用いるのが面積的に有利とされています．なお，AND ゲートを用いる場合には，乗算器の入力データが‘0’で固定されます．

実際のデータパス回路では，複数の演算器やマルチプレクサが接続され，選択制御されています．この接続関係を RTL 記述から自動で解析し，入力固定回路の挿入を行う手法が市販の EDA ツールで実用化されています．

● 低消費電力論理合成：活性化率の高い箇所の容量を低減

論理合成では通常，面積と遅延時間を小さくするような回路の最適化が行われますが，ダイナミック電力を小さくするような最適化技術も実用化されています．論理ゲートのダイナミック電力 $CV^2f\alpha$ を低減するために，スイッチング確率 α の大きい部分で負荷容量 C を減らす最適化を行います．

論理合成の過程で，スイッチング確率 α の大きいノードは論理セルの内部のノードになるようにマッピングして C を減らしたり（低電力テクノロジ・マッピング），α の大きい配線は入力容量 C の小さいピンと接続するように，論理が変わらない範囲で接続を入れ換えたり（ピン・スワップ）します．また，このほかに，クリティカル・パス以外の部分でゲートのサイズを小さくし（ゲート・サイジング），C を減らす手法も採られます．



（a）乗算器とマルチプレクサを接続した構造

（b）オペランド・アイソレーションを施した回路

図8
オペランド・アイソレーション

（a）の回路では，乗算器の出力 P が必要とされない場合でも，入力信号 B，C が変化すれば乗算器内部でスイッチングが生じ，むだな電力を消費する．オペランド・アイソレーションを施した（b）の回路では，乗算器の出力 P が必要とされるか否かをマルチプレクサの制御信号 S で検知し，必要とされない場合は乗算器の入力の値を固定する．これによって，乗算器内部でのむだなスイッチングを抑えることができる．

● マルチ V_{DD}：適材適所で二つの電源電圧を使い分ける

ダイナミック電力は $CV^2f\alpha$ で与えられることを前に述べました．この式からもわかるように，ダイナミック電力は電源電圧 V の2乗に比例します．したがって，電源電圧を下げる手法がダイナミック電力の低減化にはもっとも効果が大きいわけですが，これまで述べた手法では，電源電圧は不変として扱われてきました．電源電圧を簡単に下げられない理由は，電源電圧を下げるとMOSトランジスタの性能が低下し，回路速度が遅くなるためです．回路の性能を維持したまま消費電力を低減するには，なんらかのくふうが必要となります．

一つの方法は，回路中でほんとうに高性能が必要な部分にだけ従来の電源電圧を使い，それほど速い動作速度が必要とされない部分には低い電源電圧を使うという方法です．この手法は，低消費電力化に複数の電源電圧を使うので「マルチ V_{DD}」と呼ばれています．一般に，LSIの中でほんとうに高性能が必要とされる部分は回路全体のごく一部であり，それ以外の部分は低い電源電圧で動作させるので，全体としてダイナミック電力が小さくなります．

マルチ V_{DD} を実装する方法として，「回路ブロック単位（粗粒度）で電源電圧を変える手法」と「ゲート単位（細粒度）で電源電圧を変える手法」があり，どちらも実際のLSIで用いられています．

粗粒度のマルチ V_{DD} の例を**図9**に示します．これは米国IBM社のLSIの例(2)ですが，高速動作が要求される回路ブロック（5GHzの高速シリアル・リンク回路，**図9**の灰色で示した部分）は高い電源電圧 V_{DDH}（1.2V）で動作させ，ほかの論理ブロック（500MHz動作）およびメモリは低い電源電圧 V_{DDL}（1V）で動作させます．電源電圧の異なるブロックに信号を送出する場合には，信号の電圧振幅を変換する回路を挿入する必要があります．この変換回路はレベル・コンバータ（またはレベル・シフタ）と呼ばれます．

一方，細粒度のマルチ V_{DD} では，高速動作が必要な部分，すなわちクリティカル・パス上にあるゲートは従来の高い電圧で動作させ，クリティカル・パス上にないゲートは低電圧で動作させます．論理合成を使って生成された回路では，クリティカル・パス上のゲートは全体の中でも10～30％程度と言われており，かなりのゲートを低電圧にできます．実際の設計では，タイミング制約を満たす範囲で，低電圧で動作させるゲートをできるだけ多く見つけてダイナミック電力を小さくする手法が採られます．注意するべき点として，電源電圧の異なるゲート間で接続がある場合，レベル・コンバータをはさむ必要がある点です．レベル・コンバータは面積や電力のオーバヘッドとなるので，レベル・コンバータが必要な箇所を減らすくふうが施されます．代表的な方法として，接続しているゲートどうしはできるだけ同じ電源電圧を割り当てるようにして，論理回路を2電源化するEDA技術が開発されています．こういった細粒度のマルチ V_{DD} は，これまでに東芝(3)やIBM社(4)で社内のプロセッサに適用された事例があります．マルチ V_{DD} 手法はVoltage Island手法とも呼ばれ，近年，この手法に対応する市販EDAツールが増えてきています．

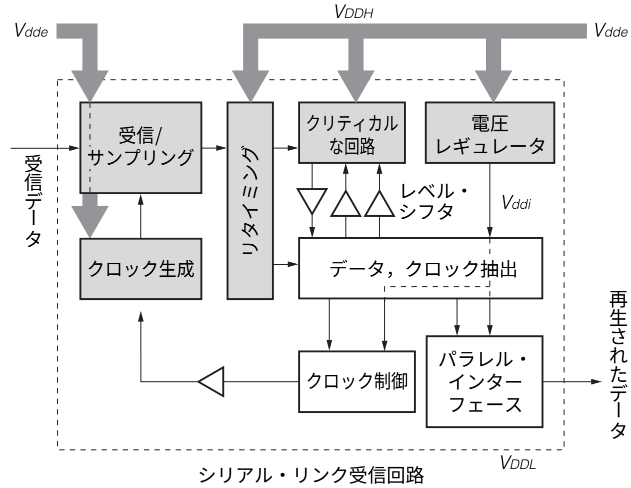


図9　粗粒度のマルチ V_{DD} を適用したチップの例

回路ブロック単位（粗粒度）で電源電圧を変える方式である．高速動作が要求される回路ブロック（5GHzの高速シリアル・リンク回路，図中の灰色で示す）は高い電源電圧 V_{DDH}（1.2V）で動作させ，ほかの論理ブロック（500MHz動作）およびメモリは低い電源電圧 V_{DDL}（1V）で動作させる．

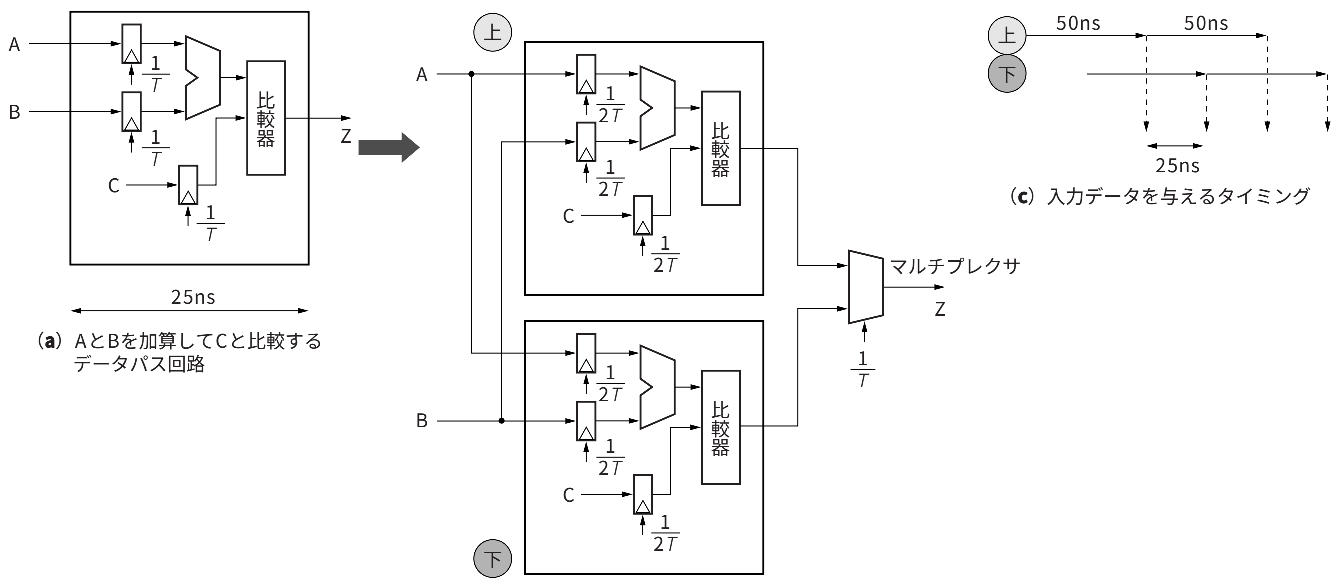
2

● 並列化を利用した低電力化：マルチコア化が進行

電源電圧を下げた場合の性能低下を，別の方法で解決する手法があります．それは，ハードウェアの並列化を行い，性能を上げたうえで低電圧化する手法です．このコンセプトは，今から15年ほど前に米国University of California, Berkeleyの研究グループが提唱し(5)，反響を呼びました．**図10**に示す例を用いて，この手法を説明します．

図10（a）に示すように，データAとBを加算し，加算結果とデータCを比較して大きいほうの値をZとして出力するデータパス回路を考えます．また，性能（スループット）に対する要求として，「25nsに1回の割合で新しいデータを出力せよ」という条件が課せられたとします．これを満





(a) AとBを加算してCと比較するデータパス回路

(b) データパスを並列化した構成

(c) 入力データを与えるタイミング

図10　並列化の利用による低電力化技術

並列化によりスループットを高めたうえで，電源電圧と動作周波数を下げて低電力化する手法である．「AとBの和とCを比較して結果をZに出力するデータパスにおいて，25nsに1回の割合で結果Zを出力せよ」という設計要求があったとする．(a)に示す構成では，25nsの遅延時間で動くデータパスが要求される．これを(b)のように，データパスを並列化したアーキテクチャに変換すると，上下のデータパスは25nsずつずらして交互に結果を出力すればよい．この方式をとると，Z出力について同じスループットを保ったまま，データパスは50nsで動作すればよいことになる．データパスに要求される速度が遅いので，電源電圧も下げられる．

たすためには，通常，A，Bのレジスタ出力からZまでの遅延時間が25nsに収まるように設計します．

一方，これと同じ性能は，**図10(b)**でも得ることができます．**図10(b)**では，**図10(a)**の加算器と比較器からなるデータパスが並列化され，しかも，上下のデータパスが25nsずつずらして交互にデータを出力するように設計されています．これらの出力データを，25nsごとにマルチプレクサで交互に切り替えてZとして出力すれば，**図10(a)**と同じスループットが得られるわけです．

ここで重要な点は，**図10(b)**では一つのデータパスは50nsの遅延時間で動作してよいという点です．**図10(c)**に示すように，上のデータパスが演算を開始してから25ns後に下のデータパスに新しい入力データを与え，さらに25ns後に上のデータパスに与え，ということを繰り返していくと，一つ一つのデータパスは50nsで動いても問題なく処理できるわけです．もともと25nsで動作するデータパスに対し，「遅延時間は50nsでよい」と許容するわけですから，このデータパスは電源電圧を下げて動作速度を落とすことができます．

前述の研究グループの評価によると(当時の電源電圧は5Vだが)，**図10(b)**では電源電圧Vを**図10(a)**の設計の0.58倍まで小さくすることができ，動作周波数fは1/2になることが示されています．一方，充放電に関係する負荷容量Cは，**図10(b)**では**図10(a)**の2.15倍に増加します．ダイナミック電力CV^2fを計算すると，結果的に，**図10(b)**は**図10(a)**の$2.15 \times 0.58^2 \times (1/2) = 0.36$倍となります．

このように，並列化によって負荷容量Cは増加するのですが，電源電圧Vが2乗で小さくなり，さらにfを下げることができるので，結果としてダイナミック電力が著しく減少します．一方，この手法ではデータパスを2個並列に設けるので，面積が大幅に増大するという問題点があり，長らく，実用レベルで広まるまでには至りませんでした．ところが，その後トランジスタの微細化が進み，1チップに搭載できる回路規模が大きくなったこと，さらに，ハイエンド(高性能)のCPUでは，動作周波数を引き上げて性能を向上させる手法では，消費電力の課題を解決できない状況を迎えたことにより，並列化を利用した低電力化の技術に，再び目が向けられるようになってきました．

近年のCPUのマルチコア化が，その現われと言えます．マルチコアでは，従来のシングル・コアと同じ演算性能を

達成するのに，より低い動作周波数で，しかも，より低い電源電圧で実現できるところが大きな利点です．これまでシングル・コアの高性能化は，動作周波数の向上に頼らざるをえない部分があったのですが，それには高性能のトランジスタを使う必要があり，リーク電流の増大を免れませんでした．

製造プロセスの微細化に伴って，リーク電力の占める割合が著しく増大しています．性能と電力の課題を満足する解として，動作周波数の向上に頼る技術には限界が来ていると言えます．並列化の利用による低電力化では，「個々のハードウェア部品は低速で動いてよい」という特徴があるので，リークの小さいトランジスタを使うことができ，リーク電力も抑えられます．同時に，低い動作周波数，低い電源電圧で実現できることから，ダイナミック電力が抑えられることは言うまでもありません．

2 リーク電力を低減する代表的な技術

リーク電力はダイナミック電力とは異なり，スイッチングを停止するだけでは低減できません．スイッチングを停止した状態でも流れ続けるリーク電流を，さまざまな技術で低減し電力削減する手法が採られます．現在のCMOS LSIではサブスレッショルド・リーク電流が支配的ですが，サブスレッショルド・リーク電流は，以下の式に示すように，トランジスタのしきい値電圧V_{th}と非常に深い関係があります．

$$I_{sub} = I_0 \exp\left(\frac{-V_{th}}{S/\ln 10}\right) \quad \cdots\cdots(2)$$

ここで，I_0は比例定数，Sは「Sファクタ」と呼ばれるパラメータで，室温では約100mV/decadeの値をとります．式(2)からわかるように，リーク電流を小さくするにはV_{th}を大きくする必要があるわけです．ところが，V_{th}はトランジスタの遅延時間t_{delay}と式(3)の関係があります．

$$t_{delay} \propto \frac{CV}{(V-V_{th})^{\alpha}} \quad \cdots\cdots(3)$$

ここで，Cは負荷容量，Vは電源電圧，αは移動度の低下を表すパラメータで，現在のデバイスでは1.1～1.5程度の値をとります．式(3)から，V_{th}を大きくすると遅延時間が増大してしまうことがわかります．このようにV_{th}を介して，遅延時間，すなわち性能とリーク電流はトレードオフの関係にあるわけです．

高性能とリーク電力低減を両立させるために，**図11**に示す技術が実用化されています．ここでは，「デュアルV_t技術」，「パワー・ゲーティング技術」，「基板バイアス制御技術」という，代表的な技術を紹介します．これらの違いは，以下の項目によって特徴をとらえることができます．

1）製造時に複数のV_{th}（マルチV_{th}）を用意するか否か

2）スリープ・モード/動作モードの切り替えを行うか否か

スリープ・モード/動作モードの切り替えを行うことで，スリープ時には低リーク電流を達成し，動作時には高速動作を行う方式として，パワー・ゲーティング技術と基板バイアス制御技術があります．パワー・ゲーティング技術の中でもポピュラなのがMTCMOS（multi threshold-voltage CMOS）と呼ばれる方式です．この方式では製造時に複数のV_{th}を用意して使用します．一方，基板バイアス制御技術では，製造時に複数のV_{th}は不要です．デュアルV_t技術は，製造時に複数のV_{th}を用意して使用する手法ですが，スリープ・モード/動作モードの切り替えはとくに行いません．以下では，これらの三つの技術について紹介します．

● デュアルV_t技術：高V_{th}と低V_{th}をチップ内で併用

高いV_{th}はリークが小さいが低速，一方の低いV_{th}はリークが大きいが高速，という性質に着目し，高V_{th}と低V_{th}を

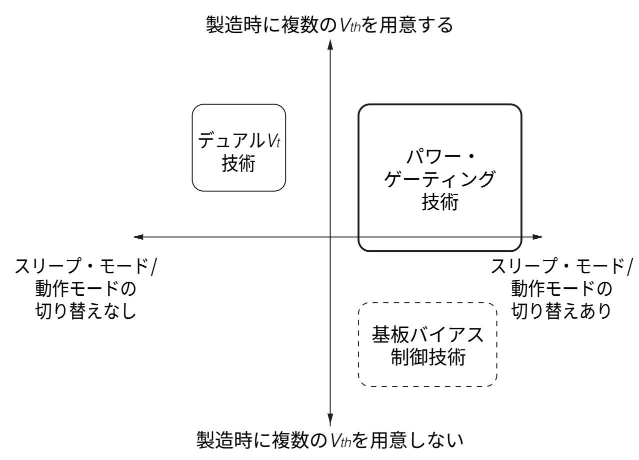
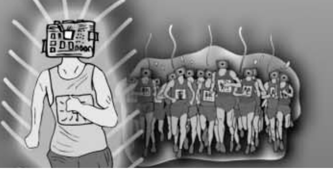


図11　リーク電力低減技術の分類

リーク電力を低減する技術では，「製造時に複数のV_{th}（マルチV_{th}）を用意するか否か」，「スリープ・モード/動作モードの切り替えを行うか否か」によって分類できる．なお，パワー・ゲーティング技術の中でも，製造時に複数のV_{th}を用意して使用する方式はMTCMOS（multi threshold-voltage CMOS）と呼ばれており，パワー・スイッチに高V_{th}を用い，論理回路部分には低V_{th}を用いる．製造時に複数のV_{th}を用意せずに，単一のV_{th}でパワー・ゲーティングを行う方式もあり，SCCMOS（super cut-off CMOS）と呼ばれている(11)．



チップ内で併用する手法が考案されました．とくに，チップ内のクリティカル・パスには低V_{th}を使い，それ以外の部分には高V_{th}を使う手法は「デュアルV_t技術」と呼ばれ，高性能なCPUチップから携帯機器用のLSIまで，非常に広い範囲で使われています．

標準的な設計手法としては，高V_{th}のセルと低V_{th}のセルをセル・ライブラリの中に用意しておき，論理合成を実行することにより，クリティカル・パス上のゲート（またはFF）には低V_{th}のセル，クリティカル・パス以外の部分には高V_{th}のセルを割り当てます．さらに別の方法として，性能を満たすために低V_{th}のセルだけを使って合成からレイアウト（配線処理）まで行い，そのあとクリティカル・パス以外の部分を抽出して，その部分だけ高V_{th}のセルと置き換える，という手法もあります．高V_{th}と低V_{th}セルでは，縦横のサイズやピン位置をまったく同一にできるので，配線後でもレイアウトを変えずにセルの置換が行えます．

最近では，半導体メーカも高V_{th}セルと低V_{th}セルを用意したセル・ライブラリを提供しており，また，デュアルV_tの設計を可能にする市販のEDAツールも続々と出てきている状況にあります．

● パワー・ゲーティング技術：スリープ時に電源を停止

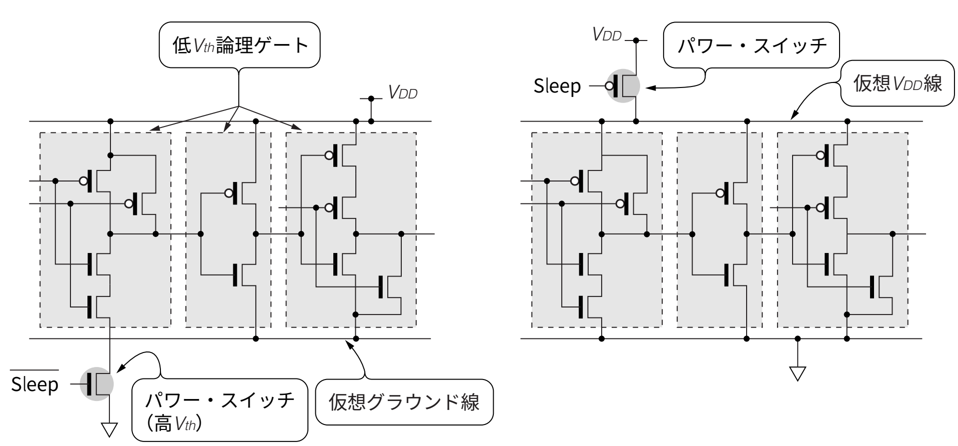
デュアルV_t技術では，クリティカル・パス上のセルは低V_{th}なので，この部分のリークは低減できていません．携帯電話などの待ち受け時など，ほとんどの回路をスリープ（スタンバイ）状態にできるときには，上記のリークをもカットできる手法が必要となります．

これを実現するのがパワー・ゲーティング技術です．パワー・ゲーティングでは，**図12（a）**に示すように，論理回路のすべてを低V_{th}のトランジスタで構成します．さらに，論理回路とグラウンドの間に，パワー・スイッチと呼ぶnMOSトランジスタを挿入します．動作時には，パワー・スイッチをONにして論理回路をグラウンドとつないで動作させます．また，スリープ時には，パワー・スイッチをOFFにして，論理回路とグラウンドを電気的にしゃ断します．高V_{th}のトランジスタをパワー・スイッチとして使う手法はもっとも典型的な手法であり，とくにMTCMOSと呼ばれます．

この手法の特徴は，動作時には低V_{th}のトランジスタで論理動作を行うため高速である点と，スリープ時には高V_{th}のトランジスタですべての論理回路のリークをしゃ断できる点です．なお，**図12（a）**のようにnMOSをパワー・スイッチとして挿入するフッタ・スイッチ方式と，**図12（b）**のようにpMOSをパワー・スイッチとして挿入するヘッダ・スイッチ方式があります．パワー・スイッチとして，V_{th}のみならずゲート絶縁膜厚t_{ox}も大きいトランジスタを使用することにより，ゲート・リークも低減することができます．

パワー・ゲーティングは，携帯電話用LSIにおける主要な技術となっていますが，パワー・ゲーティングをどういった単位で行うかについては，いくつかの方式があります．一つは，機能モジュール単位でパワー・ゲーティングを行う方式です．2006年のISSCC（International Solid-State Circuits Conference）では，チップ内部を，機能によって20個のパワー・ドメイン（電源領域）に分割し，それぞれのドメインを独立にON/OFF制御するモバイル・プロセッサLSIがルネサス テクノロジより発表(6)されました（**図13**(7)）．このチップでは，各パワー・ドメインが**図12（a）**に示す方式で構成されており，ドメインごとにパワ

図12
パワー・ゲーティング技術
パワー・スイッチに高V_{th}を用い，論理回路部分には低V_{th}を用いるMTCMOSの例である．スリープ時にはパワー・スイッチをOFFにして，電源V_{DD}からグラウンドへのリーク電流をしゃ断する．動作時にはパワー・スイッチをONにして，低V_{th}の論理ゲートで高速動作を行う．（a）のようにnMOSをパワー・スイッチとして挿入するフッタ・スイッチ方式と，（b）のようにpMOSをパワー・スイッチとして挿入するヘッダ・スイッチ方式があり，どちらも使用されている．



（a）フッタ・スイッチ方式　（b）ヘッダ・スイッチ方式

ー・スイッチの制御が行われます．

これに対して，もう少し細かい粒度でパワー・ゲーティングを行う方式も実際のチップで使われています．**図14(a)**は，セル単位でパワー・スイッチを設ける方式です．この方式の利点は，論理セルに対しパワー・スイッチが個別に付くので，論理セルの遅延時間を高い精度で予測できる点です．**図12**のようにパワー・スイッチを複数の論理ゲートで共有する方式では，論理ゲート間で放電動作が重なる場合，ゲートの出力の立ち下がり時間に影響することが知られています．一方，**図14(a)**の方式の欠点は面積オーバヘッドです．すべてのセルにパワー・スイッチを付けると面積増加が許容できなくなるので，**図14(b)**のように，回路中のクリティカル・パスだけに選択的にMTCMOSを適用する方式(セレクティブMT方式)が採られます．この手法は，東芝(8)や米国Qualcomm社(9)の携帯電話向けベースバンドLSIに適用されています．

パワー・ゲーティングを適用する部分に，FFなどの記憶回路が含まれる場合には，注意が必要です．記憶回路は，電源がしゃ断されると記憶データを保持できなくなるためです．スリープ期間にも保持が必要な記憶回路には，リテンションFFと呼ばれる特別な回路を使います．

実際に設計でパワー・ゲーティングを用いる際には，仮想グラウンド線を含む電源の配線方式の決定や，パワー・スイッチの最適サイズの決定，パワー・スイッチON時のラッシュ・カレント抑制など，さまざまな設計課題があります．パワー・ゲーティングの設計を支援するEDAツールが重要となりますが，昨年あたりから対応する市販ツールが少しずつ出始め，2006年7月に米国カリフォルニア州San Franciscoで開催されたDAC(Design Automation Conference)では，大手EDAベンダがいっせいに，パワー・ゲーティング設計を支援する機能を発表していました．

● 基板バイアス制御技術：しきい値電圧が動的に変化

基板バイアス制御技術は，基板バイアスを印加することにより，MOSトランジスタの実効しきい値を動的に変える手法です．スタンバイ時に基板に逆バイアスをかけて実効しきい値を上げる手法は，Reverse Body Biasing(RBB)と呼ばれています．当初はVTCMOS(variable threshold-voltage CMOS)(10)と呼ばれていました．この手法では，製造段階でMOSトランジスタを低いV_{th}で作っておき，動作時にはこの状態で高速論理動作を実現します．スタンバイ時には，基板バイアスを印加してトランジスタの実効V_{th}(の絶対値)を大きくし，リークを低減します．動作時とスタンバイ時で実効V_{th}を変えることにより，高速動作と低スタンバイ・リーク電流を実現する手法です．また，スタンバイ時に記憶データを保持できる点も基板バイアス制御技術の長所といえます．なお，本方式は，組み込み用マイ

2

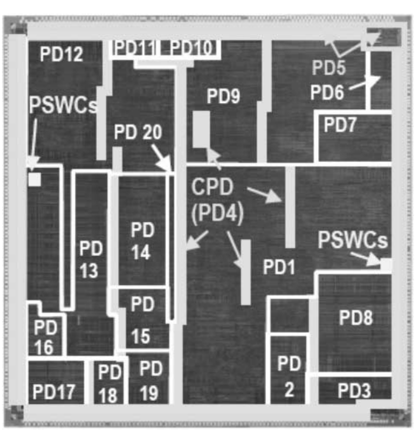
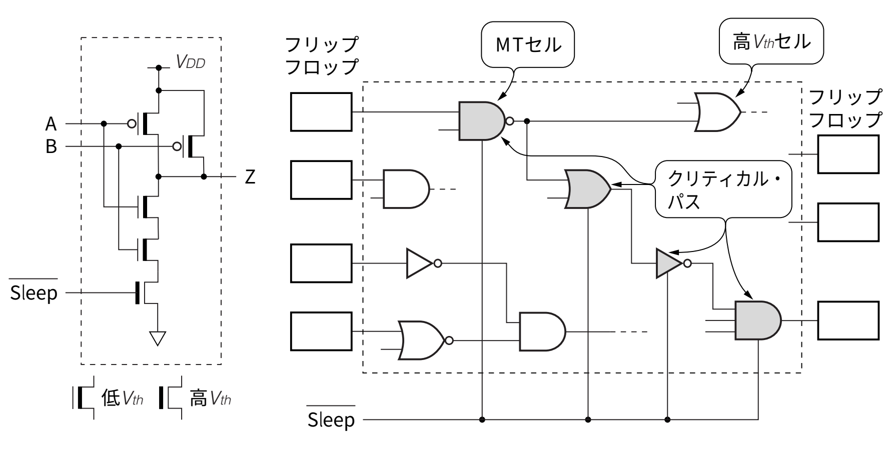


図13　チップ内部の機能モジュール単位でパワー・ゲーティングを行うモバイル・プロセッサLSI(7)

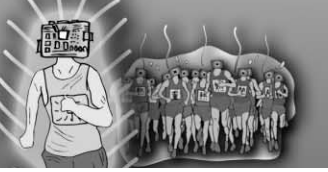
ルネサス テクノロジのモバイル・プロセッサLSIでは，チップ内部を機能によって20個のパワー・ドメイン(電源領域)に分割し，パワー・ゲーティングを用いてそれぞれのドメインを独立にON/OFF制御する．例えば，テレビ電話を実現するときには，ほぼすべてのパワー・ドメインをONさせるが，電話の待ち受け時には，パワー・ドメインPD13やPD14などの限られたドメイン以外はすべてパワー・スイッチをOFFにして，リークをしゃ断する．



(a) MTセル(2入力NAND)　　(b) セレクティブMT方式

図14　セル単位でパワー・ゲーティングを行う方式

細粒度でパワー・ゲーティングを行う方式として，(a)のようにセル単位でパワー・スイッチを設ける方式も使われている．セル内部にパワー・スイッチを持たせたセルはMT(multi-threshold)セルと呼ばれ，(b)に示すセレクティブMT方式では，クリティカル・パスのみに選択的にMTセルを適用する．非クリティカル・パスには高V_{th}セルを用いる．



クロプロセッサのXScaleなどで使用されています[11]．

一方，基板に逆バイアスをかける上記の方式について，トランジスタの微細化を進めていく際の問題点が指摘されています．微細化を進めるにつれて，サブスレッショルド・リークを最小化するには，これまでより大きな逆バイアスの値が必要となるためです．逆バイアスの値を大きくすると，バンド間トンネリングという効果のため，ジャンクション・リークが増大してしまい，リーク削減効果が薄れる結果となってしまいます．リーク低減効果を維持するには，トランジスタの製造技術の改良が必要となります．

一方で，このような逆バイアスを印加する方式ではなく，順バイアスを印加するForward Body Biasing（FBB）という手法も研究されています．この方式では，製造段階において高いV_{th}でトランジスタを作っておき，スタンバイ時にはこの状態で低リークを達成します．動作時にはわずかな順バイアスを印加して実効しきい値を低くし，高速論理動作を実現します．FBB方式には微細化に対しても効果を維持できるという性質があり，米国Intel社で研究開発が進められています．

● 今後，動作時のリーク電力が問題に

集積回路の低消費電力技術として，ダイナミック電力とリーク電力を低減する代表的な設計技術を紹介しました．

これまではどちらかというと，「動作時にはダイナミック電力が重要で，スタンバイ時にはリーク電力が重要である」という構図がありました．しかし，半導体製造技術の微細化に伴うリーク電流の増大により，この構図は変わっていく傾向にあります．すなわち，動作時においてもリーク電力の低減化を真剣に考えなければならなくなります．

今後，この分野の新しい技術が重要性を増していくと思われます．

参考・引用*文献

(1) R. Puri, et al ; " Keeping Hot Chips Cool ", *Proceedings of Design Automation Conference*（*DAC* '05）, pp.285-288, June 2005.

(2) J. Carballo, et al ; " A Semi-Custom Voltage-Island Technique and Its Application to High-Speed Serial Links ", *Proceedings of International Symposium on Low Power Electronics and Design*（*ISLPED* '03）, pp.60-65, Aug. 2003.

(3) K. Usami, et al ; " Design Methodology of Ultra Low-power MPEG4 Codec Core Exploiting Voltage Scaling Techniques ", *Proceedings of Design Automation Conference*（*DAC* '98）, pp.483-488, June 1998.

(4) R. Puri, et al ; " Pushing ASIC Performance in a Power Envelope ", *Proceedings of Design Automation Conference*（*DAC* '03）, pp. 788-793, June 2003.

(5) A. Chandrakasan, et al ; " Low-power CMOS digital design ", *IEEE Journal of Solid-State Circuits*, vol.27, no.4, April 1992.

(6) T. Hattori, et al ; " A Power Management Scheme Controlling 20 Power Domains for a Single-Chip Mobile Processor ", *ISSCC* 2006 *Digest of Technical Papers*, pp. 542-543, Feb. 2006.

(7) T. Hattori, et al ; " Hierarchical Power Distribution and Power Management Scheme for a Single Chip Mobile Processor ", *Proceedings of Design Automation Conference*（*DAC* '06）, pp. 292-295, July 2006.

(8) K. Usami, et al ; " Automated Selective Multi-Threshold Design for Ultra-Low Standby Applications ", *Proceedings of International Symposium on Low Power Electronics and Design*（*ISLPED* '02）, pp.202-206, Aug. 2002.

(9) G. Uvieghara, et al ; " A Highly-Integrated 3G CDMA2000 1X Cellular Baseband Chip with GSM/AMPS/GPS/Bluetooth/Multimedia Capabilities and ZIF RF Support ", *ISSCC* 2004 *Digest of Technical Papers*, pp. 422-536, Feb. 2004.

(10) T. Kuroda, et al ; " A 0.9V, 150MHz, 10mW, 4mm^2, 2-D Discrete Cosine Transform Core Processor with Variable Threshold-Voltage（VT）Scheme ", *IEEE Journal of Solid-State Circuits*, vol.31, no.11, pp. 1770-1779, Nov. 1996.

(11) S. Narendra and A. Chandrakasan（ed.）; *Leakage in Nanometer CMOS Technologies*, Springer, 2006.

うさみ・きみよし
芝浦工業大学 工学部情報工学科 教授

<筆者プロフィール>
宇佐美公良．1982年，早稲田大学 理工学部電気工学科卒．1984年，早稲田大学大学院 理工学研究科修士過程修了．同年，東芝入社．半導体技術研究所にてマイクロプロセッサの開発設計に従事．1993～1995年，米国Stanford University客員研究員．低消費電力設計技術に関する研究を始める．東芝に帰任後，LSIの低電力設計技術，CAD技術の開発と事業展開に従事．2000年，早稲田大学より博士（工学）を授与．2003年より芝浦工業大学 工学部情報工学科助教授．2005年より同大学教授．研究分野は，システムLSIの低消費電力設計技術．